REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggesstions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any oenalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
13-01-2013	Conference Proceeding	-

4. TITLE AND SUBTITLE

DISCRIMINATIVE GRAPHICAL MODELS FORSPARSITY-BASED HYPERSPECTRAL TARGET DETECTION

5a. CONTRACT NUMBER
W911NF-11-1-0245

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

Umamahesh Srinivas,, Yi Chen,, Vishal Monga,, Nasser M. Nasrabadi,, Trac D. Tran

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

Johns Hopkins University
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218 -2686

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
60291-MA.6

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

The inherent discriminative capability of sparse representations has been exploited recently for hyperspectral target detection. This approach relies on the observation that the spectral signature of a pixel can be represented as a linear combination of a few training spectra drawn from both

15. SUBJECT TERMS

Hyperspectral target detection, sparsity,probabilistic graphical models

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Trac Tran
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 410-516-7416

Report Title

DISCRIMINATIVE GRAPHICAL MODELS FORSPARSITY-BASED HYPERSPECTRAL TARGET DETECTION

ABSTRACT

The inherent discriminative capability of sparse representations has been exploited recently for hyperspectral target detection. This approach relies on the observation that the spectral signature of a pixel can be represented as a linear combination of a few training spectra drawn from both target and background classes. The sparse representation corresponding to a given test spectrum captures class-specific discriminative information crucial for detection tasks. Spatiospectral information has also been introduced into this framework via a joint sparsity model that simultaneously solves for the sparse features for a group of spatially local pixels, since such pixels are highly likely to have similar spectral characteristics. In this paper, we propose a probabilistic graphical model framework that can explicitly learn the class conditional correlations between these distinct sparse representations corresponding to different pixels in a spatial neighborhood. Simulation results show that the proposed algorithm outperforms classical hyperspectral target detection algorithms as well as support vector machines.

Conference Name: IEEE International Geoscience and Remote Sensing Symposium (IGARSS2012)

Conference Date: July 22, 2012

DISCRIMINATIVE GRAPHICAL MODELS FOR SPARSITY-BASED HYPERSPECTRAL TARGET DETECTION

Umamahesh Srinivas[†], *Yi Chen*[‡], *Vishal Monga*[†], *Nasser M. Nasrabadi*[§], *and Trac D. Tran*[‡]

†Dept. of Electrical Engineering, Pennsylvania State University, University Park, PA, USA
‡Dept. of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD, USA
§US Army Research Laboratory, Adelphi, MD, USA

ABSTRACT

The inherent discriminative capability of sparse representations has been exploited recently for hyperspectral target detection. This approach relies on the observation that the spectral signature of a pixel can be represented as a linear combination of a few training spectra drawn from both target and background classes. The sparse representation corresponding to a given test spectrum captures class-specific discriminative information crucial for detection tasks. Spatio-spectral information has also been introduced into this framework via a joint sparsity model that simultaneously solves for the sparse features for a group of spatially local pixels, since such pixels are highly likely to have similar spectral characteristics. In this paper, we propose a probabilistic graphical model framework that can explicitly learn the *class conditional correlations* between these distinct sparse representations corresponding to different pixels in a spatial neighborhood. Simulation results show that the proposed algorithm outperforms classical hyperspectral target detection algorithms as well as support vector machines.

Index Terms— Hyperspectral target detection, sparsity, probabilistic graphical models.

1. INTRODUCTION

An important research problem in hyperspectral imaging (HSI) [1] is hyperspectral target detection, which can be viewed as a binary classification problem where hyperspectral pixels are labeled as either target or background based on their spectral characteristics. Many statistical hypothesis testing techniques [2] have been proposed for hyperspectral target detection, including the spectral matched filter (SMF), matched subspace detector (MSD) and adaptive subspace detector (ASD). Advances in machine learning theory have contributed to the popularity of support vector machines (SVM) [3] as a powerful tool to classify hyperspectral data.

A significant recent advance has exploited the inherent discriminative nature of sparse representations for hyperspectral target detection [4]. The sparsity model is posited on the observation that spectral signatures of pixels from the same class (target or background) lie in a low-dimensional subspace. Consequently, the spectral signature of a test pixel can be represented by the linear combination of a few training spectra which come from an over-complete dictionary built using the target and background subspaces. The associated sparse representation, which is obtained as the solution to a sparsity-constrained optimization problem, has been shown to capture class-specific discriminative information crucial for detection and classification tasks. Typically in hyperspectral images, pixels in a small spatial neighborhood belong to the same class and their spectra are highly correlated, a fact not exploited by the pixel-wise sparse representation model. To address this issue, a joint sparsity model has been proposed recently [5] to simultaneously capture spatial and spectral characteristics. The spectral signatures of pixels in a local spatial neighborhood (of the pixel of interest) are constrained to be represented by a *common* collection of training spectra, albeit with different weights. A simultaneous sparse recovery problem is now solved to recover both the training support and the corresponding sparse representation vectors.

Motivation: The resulting sparse representations are discriminative in nature. Hence, the detection statistic in [4,5] involves a simple comparison of reconstruction residuals using the training and background subspaces separately. The sparse representations corresponding to different pixels in a local neighborhood are statistically correlated, and this correlation is captured intuitively by the joint sparsity model. A challenging open problem, therefore, is to mine the *class-conditional correlations* among these distinct feature representations in a more principled manner for detection and classification. In this paper, we propose a probabilistic graphical model framework to explicitly learn the conditional dependencies between the sparse features via discriminative graphs.

Overview of contribution: The sparse representation vectors of pixels at different locations in a local spatial neighborhood (relative to a central pixel) comprise several distinct sets of features which provide complementary yet correlated information useful for detection. To learn these

This work has been partially supported by NSF under Grants CCF-1117545 and ARO under Grant 60219-MA.

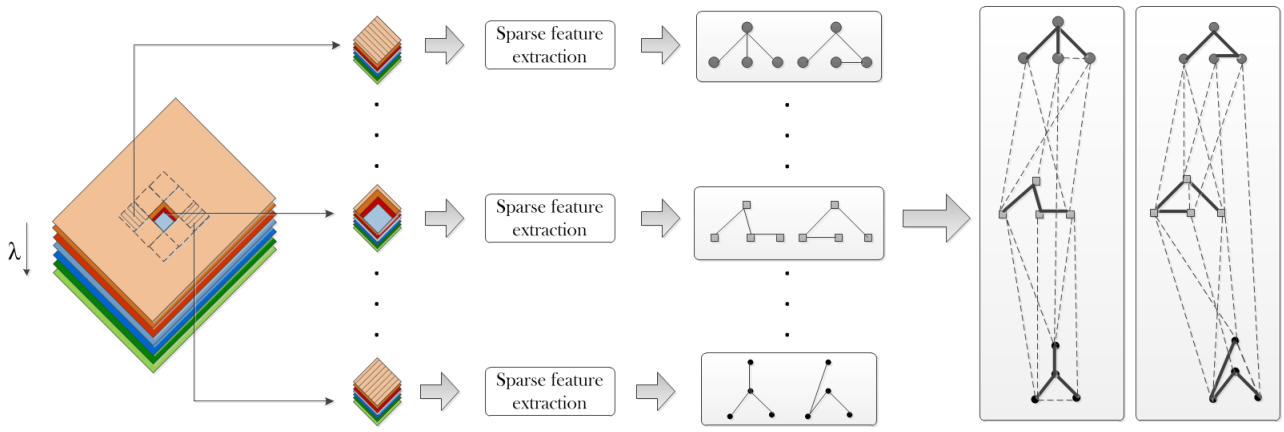

Fig. 1. Hyperspectral image detection using discriminative graphical models on sparse feature representations obtained from local pixel neighborhoods. The rightmost figure shows the final learnt graphs, where solid lines represent the initial disjoint graphs and the dashed lines represent newly learnt edges which capture conditional correlations.

statistical correlations, we first learn a pair of *discriminative tree graphs* (one for each class) for each distinct set of features [6], and then augment new edges to these initially disjoint graphs iteratively via boosting [7]. Consequently, we learn a discriminative classifier over the sparse features unlike the reconstruction residual-based detection scheme in [5].

2. BACKGROUND

2.1. Sparsity Models for Hyperspectral Target Detection

Let $\boldsymbol{y} \in \mathbb{R}^B$ be a pixel with B indicating the number of spectral bands, $\boldsymbol{D}_b \in \mathbb{R}^{B \times N_b}$ be the sub-dictionary whose columns are the N_b background training samples, and $\boldsymbol{D}_t \in \mathbb{R}^{B \times N_t}$ be the sub-dictionary whose columns are the N_t background training samples. The HSI pixel $\boldsymbol{y}$ can then be written as:

$$\boldsymbol{y} = \boldsymbol{D}_b \boldsymbol{\alpha}_b + \boldsymbol{D}_t \boldsymbol{\alpha}_t = \underbrace{\begin{bmatrix} \boldsymbol{D}_b & \boldsymbol{D}_t \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\begin{bmatrix} \boldsymbol{\alpha}_b \\ \boldsymbol{\alpha}_t \end{bmatrix}}_{\boldsymbol{\alpha}} = \boldsymbol{D}\boldsymbol{\alpha}, \tag{1}$$

where $\boldsymbol{D} \in \mathbb{R}^{B \times N}$ with $N = N_b + N_t$ is a dictionary consisting of training samples from both training and background classes, and $\boldsymbol{\alpha} \in \mathbb{R}^N$ is a sparse vector. Given the overcomplete dictionary $\boldsymbol{D}$, the sparse coefficient vector $\boldsymbol{\alpha}$ is obtained by solving the following optimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{\alpha}\|_2 \leq \varepsilon, \tag{2}$$

where ε is a suitably chosen reconstruction error tolerance. The class label of $\boldsymbol{y}$ is determined by comparing the reconstruction residuals:

$$R(\boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{D}_b \hat{\boldsymbol{\alpha}}_b\|_2 - \|\boldsymbol{y} - \boldsymbol{D}_t \hat{\boldsymbol{\alpha}}_t\|_2, \tag{3}$$

where $\hat{\boldsymbol{\alpha}}_b$ and $\hat{\boldsymbol{\alpha}}_t$ are, respectively, the set of coefficients in $\hat{\boldsymbol{\alpha}}$ corresponding to $\boldsymbol{D}_b$ and $\boldsymbol{D}_t$. The test vector $\boldsymbol{y}$ is identified as a target pixel if $R(\boldsymbol{y})$ is larger than some suitably chosen positive threshold δ; if not, it is labeled as a background pixel.

This pixel-wise sparsity model is extended to incorporate local spatial information in [5] by enforcing a common support set of training spectra for a collection of neighboring pixels $\boldsymbol{y}_i, i = 1, \ldots, T$, as follows:

$$\begin{aligned} \boldsymbol{Y} &= \begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_T \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}\boldsymbol{\alpha}_1 & \boldsymbol{D}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{D}\boldsymbol{\alpha}_T \end{bmatrix} \\ &= \boldsymbol{D} \underbrace{\begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_T \end{bmatrix}}_{\boldsymbol{S}} = \boldsymbol{D}\boldsymbol{S}. \end{aligned} \tag{4}$$

Since all the pixels in $\boldsymbol{Y}$ are represented by the same collection of training spectra in $\boldsymbol{D}$, the vectors $\boldsymbol{\alpha}_i, i = 1, \ldots, T$, all have non-zero entries at the same locations. As a result, $\boldsymbol{S}$ is a sparse matrix with only a few nonzero rows, and is recovered by solving the following constrained optimization problem:

$$\hat{\boldsymbol{S}} = \arg\min \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{S}\|_F \quad \text{subject to} \quad \|\boldsymbol{S}\|_{\text{row},0} \leq K_0, \tag{5}$$

where $\|\boldsymbol{S}\|_{\text{row},0}$ denotes the number of non-zero rows of $\boldsymbol{S}$ and $\|\cdot\|_F$ is the Frobenius norm. The problem in (5) can be approximately solved by the greedy Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm [8].

2.2. Probabilistic Graphical Models

Probabilistic graphical models provide a convenient way of visualizing the correlations between the individual random variables in a multivariate probability distribution. The random variables are represented by the nodes $\mathcal{V} = \{v_1, \ldots, v_r\}$ in a graph $\mathcal{G}$, and the (undirected) edges $\mathcal{E} \subset \binom{\mathcal{V}}{2}$ which connect pairs of nodes identify conditional dependencies. A graphical model hence approximates the joint probability distribution function by a product of terms that represent pairwise and marginal statistics. Many recent applications [9] have demonstrated the ability of graphical models to learn

models for high-dimensional data using limited training (a typical scenario for practical HSI applications) under moderate computational complexity.

As the starting point for our contribution, we consider a recent discriminative learning framework [6] wherein a pair of graphs is jointly learnt by minimizing the classification error. Specifically, the tree-approximate J-divergence (a symmetric extension of the Kullback-Leibler (KL) distance) between two distributions p and q is maximized:

$$\hat{J}(\hat{p},\hat{q};p,q) = \int (p(x)-q(x)) \log \left[\frac{\hat{p}(x)}{\hat{q}(x)} \right] dx. \qquad (6)$$

Based on the observation that maximizing the J-divergence minimizes the upper bound on the probability of classification error, the discriminative learning problem then becomes:

$$(\hat{p},\hat{q}) = \arg \max_{\hat{p},\hat{q} \text{ are trees}} \hat{J}(\hat{p},\hat{q};\tilde{p},\tilde{q}), \qquad (7)$$

where $\tilde{p}$ and $\tilde{q}$ are the available empirical estimates. The problem in (7) is shown to decouple into two maximum-weight spanning tree (MWST) problems [6]:

$$\begin{aligned} \hat{p} &= \arg \min_{\hat{p} \text{ is a tree}} D(\tilde{p}||\hat{p}) - D(\tilde{q}||\hat{p}) \\ \hat{q} &= \arg \min_{\hat{q} \text{ is a tree}} D(\tilde{q}||\hat{q}) - D(\tilde{p}||\hat{q}), \end{aligned} \qquad (8)$$

where $D(p||\hat{p}) = E_p[\log(p/\hat{p})]$ represents the KL-distance. From (8), we see that the optimal choice of $\hat{p}$ ($\hat{q}$) minimizes its distance to $\tilde{p}$ ($\tilde{q}$) while simultaneously maximizing its distance from $\tilde{q}$ ($\tilde{p}$). The trade-off between generalization and performance inherent to graphical models is resolved by iteratively thickening the initial graph with more edges via boosting [7] to learn a richer structure.

As discussed earlier, the sparse representations from different pixels in a local spatial neighborhood are correlated and our contribution is an attempt to explicitly learn these conditional dependencies. To this end, we instantiate our recent discriminative graphical framework [10] for HSI detection.

3. DISCRIMINATIVE GRAPHICAL MODELS FOR HYPERSPECTRAL TARGET DETECTION

In this section, we introduce our proposed Local-Sparsity-Graphical-Model (LSGM) approach for joint sparsity and graphical model-based HSI detection. An illustration of the overall framework is shown in Fig. 1. Algorithm 1 outlines the steps in the process, which consists of an offline training stage (Steps 1-4) followed by an online test stage (Steps 5-6). The discriminative graphs are learnt in the training stage. First, feature vectors (i.e., sparse vectors with respect to a given $\boldsymbol{D}$) of training samples and their neighboring pixels are obtained by solving the joint sparse recovery problem in (5).

Let T be the size of the neighborhood. For every pixel $\boldsymbol{y} \in \mathbb{R}^B$, T different features $\boldsymbol{\alpha}_l \in \mathbb{R}^N, l = 1,2,\ldots,T$ are obtained, as illustrated in Fig. 1 for a 3×3 neighborhood with $T = 9$. Training features for class C_t correspond to pixels in a neighborhood of training target samples, while features for C_b are the sparse vectors associated with neighbors of background training samples. For each of the T sets of features, a pair of N-node discriminative tree graphs $\mathcal{G}_l^t$ and $\mathcal{G}_l^b$, which respectively approximate the class distributions $f(\boldsymbol{\alpha}_l|C_t)$ and $f(\boldsymbol{\alpha}_l|C_b)$, are simultaneously learnt. The initial disjoint graphs with TN nodes representing the class distribution corresponding to C_t and C_b are then generated by separately concatenating the nodes of $\mathcal{G}_l^t, l = 1,\ldots,T$ and $\mathcal{G}_l^b, l = 1,\ldots,T$, respectively. These graphs with sparse edge structure are then iteratively thickened via boosting [10]. Different pairs of discriminative graphs over the same sets of nodes with different weights are learnt in different iterations, and the newly-learnt edges are used to augment the graphs. The final "thickened" graphs $\mathcal{G}^t$ and $\mathcal{G}^b$ are shown in Fig. 1.

Algorithm 1 LSGM (Steps 1-4 offline)

1: **Feature extraction (training):** Compute sparse representations $\boldsymbol{\alpha}_l, l = 1,\ldots,T$ for neighboring pixels of the training data
2: **Initial disjoint graphs:** Discriminatively learn T pairs of N-node tree graphs $\mathcal{G}_l^t$ and $\mathcal{G}_l^b$ on $\{\boldsymbol{\alpha}_l\}$, for $l = 1,\ldots,T$, obtained from training data
3: Separately concatenate nodes corresponding to the two classes, to generate initial graphs
4: **Boosting on disjoint graphs**: Iteratively thicken initial disjoint graphs via boosting to obtain final graphs $\mathcal{G}^t$ and $\mathcal{G}^b$

{**Online process**}

5: **Feature extraction (test):** Obtain sparse representations $\boldsymbol{\alpha}_l, l = 1,\ldots,T$ in $\mathbb{R}^N$ from test image
6: **Inference**: Classify based on output of the resulting classifier using (9).

The above process is performed offline. The classification of a new test sample is then performed online. Features $\boldsymbol{\alpha}$ are extracted from the test sample $\boldsymbol{y}$ by solving the sparse recovery problem in (5) for the T pixels in the neighborhood centered at $\boldsymbol{y}$. Let $\hat{f}(\boldsymbol{\alpha}|C_t)$ and $\hat{f}(\boldsymbol{\alpha}|C_b)$ denote the probability distribution functions for the final graphs $\mathcal{G}^t$ and $\mathcal{G}^b$ learnt for C_t and C_b respectively. The class label of $\boldsymbol{y}$ is finally determined as follows:

$$\text{Class}(\boldsymbol{y}) = \begin{cases} \text{Target} & \text{if } \log\left(\frac{\hat{f}(\boldsymbol{\alpha}|C_t)}{\hat{f}(\boldsymbol{\alpha}|C_b)}\right) \geq 0 \\ \text{Background} & \text{if } \log\left(\frac{\hat{f}(\boldsymbol{\alpha}|C_t)}{\hat{f}(\boldsymbol{\alpha}|C_b)}\right) < 0. \end{cases} \qquad (9)$$

4. EXPERIMENTAL RESULTS AND DISCUSSION

Hyperspectral images from the HYDICE forest radiance I data collection (FR-I) [11] are used for the experiment. The HYDICE sensor generates 210 bands across the whole spectral range from 0.4 to 2.5 μm, spanning the visible and short-wave infrared bands and including 14 targets. Only 150 of the 210 available bands are retained by removing the absorption and low-SNR bands. The target sub-dictionary $\boldsymbol{D}_t$ comprises 18 training spectra chosen from the leftmost target

Table 1. Confusion matrix for the FR-I hyperspectral image. Four different methods are compared. ($N_t = 18$ and $N_b = 216$.)

Class	Target	Background	Method
Target	0.6512	0.3488	MSD
	0.9493	0.0507	SVM-CK
	0.9556	0.0444	SOMP
	0.9612	**0.0388**	**LSGM**
Background	0.0239	0.9761	MSD
	0.0090	0.9910	SVM-CK
	0.0097	0.9903	SOMP
	0.0086	**0.9914**	**LSGM**

in the scene, while the background sub-dictionary $\boldsymbol{D}_b$ has 216 training spectra chosen using the dual window technique described in [4].

Four different methods are compared: (i) classical matched subspace detector (MSD) which operates on each pixel independently [12], (ii) composite kernel support vector machines (SVM-CK) which considers a weighted sum of spectral and spatial information [3], (iii) simultaneous orthogonal matching pursuit (SOMP) which involves solving Eq. (5) with a 3×3 local window [5], and (iv) the proposed LSGM approach with the same 3×3 window to generate the sparse features. Table 1 shows the confusion matrix in which detection and error rates are provided with each row representing the true class of the test pixels and each column representing the output of the specified classifier. All four approaches are compared, and the proposed LSGM methods offers better target detection performance. Improvements over SOMP can be attributed to the use of an explicit discriminative classifier in LSGM. All approaches identify the background class with a reasonably high degree of accuracy.

Fig. 2 shows the receiver operating characteristics (ROC) curve for the detection problem. The ROC curve describes the probability of detection (PD) as a function of the probability of false alarms (PFA). To calculate the ROC curve, a large number of thresholds are chosen between the minimum and maximum of the detector output, and class labels for all test pixels are determined at each threshold. The PFA is calculated as the ratio of the number of false alarms (background pixels determined as target) to the total number of pixels in the test region, while the PD is the ratio of the number of hits (target pixels correctly determined as target) to the total number of true target pixels. It can be seen that the proposed LSGM approach offers the best overall detection performance.

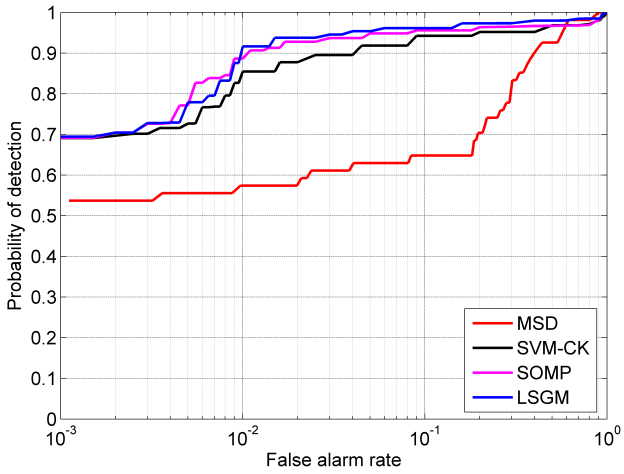

Fig. 2. ROC for FR-I comparing the four approaches: (i) MSD [12], (ii) SVM-CK [3], (iii) SOMP [5], and (iv) the proposed algorithm (LSGM).

5. REFERENCES

[1] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral Remote Sensing - Principles and Applications*, CRC Press, Boca Raton, FL, USA, 2008.

[2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[3] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sensing Letters*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

[4] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 629–640, June 2011.

[5] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geosci. Remote Sensing Letters*, vol. 8, no. 4, pp. 676–680, July 2011.

[6] V. Y. F. Tan, S. Sanghavi, J. W. Fisher, and A. S. Willsky, "Learning graphical models for hypothesis testing and classification," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5481–5495, Nov. 2010.

[7] Y. Freund and R. E. Schapire, "A short introduction to boosting," *Journal of Japanese Society for Artificial Intelligence*, vol. 14, no. 5, pp. 771–780, Sept. 1999.

[8] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, pp. 572–588, Mar. 2006.

[9] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families and variational inference," *Foundations and Trends in Machine Learning*, vol. 1, no. 1-2, pp. 1–305, 2008.

[10] U. Srinivas, V. Monga, and R. G. Raj, "Automatic target recognition using discriminative graphical models," in *Proc. IEEE Intl. Conf. Image Processing*, 2011, pp. 33–36.

[11] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "HYDICE system: Implementation and performance," in *Proc. SPIE Conf. Algorithms Technol. Multispectral, Hyperspectral, Ultraspectral Imagery XV*, 1995, pp. 258–267.

[12] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.